MBMG
Metabarcoding & Metagenomics

Metabarcoding and Metagenomics 4: 151–163
DOI 10.3897/mbmg.4.58572

Monitoring Schema

Methodological considerations for monitoring soil/litter arthropods in tropical rainforests using DNA metabarcoding, with a special emphasis on ants, springtails and termites

Yves Basset[1,2,3,4], **David A. Donoso**[5], **Mehrdad Hajibabaei**[6], **Michael T. G. Wright**[6], **Kate H. J. Perez**[6], **Greg P. A. Lamarre**[1,2,3], **Luis F. De León**[7,8], **José G. Palacios-Vargas**[9], **Gabriela Castaño-Meneses**[10], **Marleny Rivera**[1], **Filonila Perez**[1], **Ricardo Bobadilla**[1], **Yacksecari Lopez**[1], **José Alejandro Ramirez**[1], **Héctor Barrios**[4]

1 *ForestGEO, Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, Ancon, Panamá*
2 *Faculty of Science, University of South Bohemia, 370 05, Ceske Budejovice, Czech Republic*
3 *Biology Centre of the Czech Academy of Sciences, Institute of Entomology, 370 05, Ceske Budejovice, Czech Republic*
4 *Maestría de Entomología, Universidad de Panamá, 080814, Panama City, Republic of Panama*
5 *Departamento de Biología, Escuela Politécnica Nacional, Quito, Ecuador*
6 *Centre for Biodiversity Genomics, 50 Stone Road East, University of Guelph, Guelph, ON, N1G2W1, Canada*
7 *Department of Biology, Integrated Science Complex (ISC 3-3100), University of Massachusetts Boston, 100 Morrissey Blvd., Boston, MA 02125, USA*
8 *Centro de Biodiversidad y Descubrimiento de Drogas, Instituto de Investigaciones Científicas y Servicios de Alta Tecnología (INDICASAT-AIP), Panamá City, Panamá*
9 *Laboratorio de Ecología y Sistemática de Microartrópodos, Departamento de Ecología y Recursos Naturales. Facultad de Ciencias, Universidad Nacional Autónoma de México, 04510, Mexico City, Mexico*
10 *Unidad Multidisciplinaria de Docencia e Investigación, Facultad de Ciencias, Universidad Nacional Autónoma de México, Campus Juriquilla. Juriquilla 76230, Querétaro, Mexico*

Corresponding author: Yves Basset (bassety@si.edu)

Academic editor: Florian Leese | **Received** 15 September 2020 | **Accepted** 10 December 2020 | **Published** 31 December 2020

Abstract

Robust data to refute or support claims of global insect decline are currently lacking, particularly for the soil fauna in the tropics. DNA metabarcoding represents a powerful approach for rigorous spatial and temporal monitoring of the taxonomically challenging soil fauna. Here, we provide a detailed field protocol, which was successfully applied in Barro Colorado Island (BCI) in Panama, to collect soil samples and arthropods in a tropical rainforest, to be later processed with metabarcoding. We also estimate the proportion of soil/litter ant, springtail and termite species from the local fauna that can be detected by metabarcoding samples obtained either from Berlese-Tullgren (soil samples), Malaise or light traps. Each collecting method detected a rather distinct fauna. Soil and Malaise trap samples detected 213 species (73%) of all target species. Malaise trap samples detected many ant species, whereas soil samples were more efficient at detecting springtail and termite species. With respect to long-term monitoring of soil-dwelling and common species (more amenable to statistical trends), the best combination of two methods were soil and light trap samples, detecting 94% of the total of common species. A protocol including 100 soil, 40 Malaise and 80 light trap samples annually processed by metabarcoding would allow the long-term monitoring of at least 11%, 18% and 16% of species of soil/litter ants, springtails and termites, respectively, present on BCI, and a high proportion of the total abundance (up to 80% of all individuals) represented by these taxa.

Key Words

Barro Colorado Island, Berlese-Tullgren, Collembola, Formicidae, Isoptera, light trap, Malaise trap

Copyright *Yves Basset et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

PENSOFT

Introduction

Arthropods represent the majority of macroscopic terrestrial animal life, both in terms of species richness (Stork 2018) and biomass (Pimentel and Andow 1984), particularly in the tropics (Basset et al. 2012). They also provide critical services, including ecosystem functioning and human food security. Despite their ecological and economic importance, arthropods are still comparatively neglected as a primary focus of scientific research, limiting our understanding of the functional ecology of ecosystems (Basset et al. 2019). Alarmingly, recent reports suggest catastrophic declines in current insect abundance, with potentially serious implications for ecosystem functioning. This includes declines in the richness of insect species in temperate countries (Sánchez-Bayo and Wyckhuys 2019), insect biomass in Germany and Puerto Rico (Hallmann et al. 2017; Lister and Garcia 2018), butterflies and moths in the United Kingdom (McDermott Long et al. 2017), or insect pollinators worldwide (Goulson et al. 2015). Although some of these articles have attracted considerable attention in the media (often with dramatic epithets such as "Insect Armageddon"), they have also been criticized for lack of scientific rigor (Leather 2017; Didham et al. 2020). What is clear is that we currently lack robust data to refute or support claims of global insect decline, particularly in the tropics (Basset and Lamarre 2019; Crossley et al. 2020; van Klink et al. 2020; Wagner 2020), not to mention scenarios of how insect populations will be affected by anthropogenic changes in the future (Lamarre et al. 2020).

In the tropics, especially in tropical rainforests, arthropods may face significant threats due to habitat loss (Wagner 2020), as well as climate change (Deutsch et al. 2008). High quality data on the population dynamics of tropical insects are urgently needed to understand the implications of insect decline (Basset and Lamarre 2019; Crossley et al. 2020; Wagner 2020). This appears even more important for the soil fauna in the tropics, which remains poorly surveyed, known and understood (André et al. 2002). For example, out of the 17 insect monitoring studies performed in the tropics with time-series > 5 years reviewed by Basset and Lamarre (2019), only one study targeted the soil fauna (litter ants: Donoso 2017). To the best of our knowledge, no long-term monitoring program is currently focused on the soil fauna in the tropics.

DNA metabarcoding (Shendure and Ji 2008; hereafter "metabarcoding") represents a powerful approach for the screening of numerous environmental samples rich in species and for rigorous spatial and temporal monitoring (Leray and Knowlton 2015; Tang et al. 2015; Beng et al. 2016). In recent years, metabarcoding has proven useful for comparing local soil communities (Arribas et al. 2016, Oliverio et al. 2018), including in tropical rainforests (Zinger et al. 2019).

A rare effort to monitor arthropods in tropical rainforests in the long term is epitomized by the ForestGEO Arthropod Initiative in Panama (Lamarre et al. 2020). Arthropod surveys are performed within permanent forest dynamics plots monitored by the Forest Global Earth Observatories (ForestGEO; http://www.forestgeo.si.edu/; Anderson-Teixeira et al. 2015). In Panama, the ForestGEO Arthropod Initiative has been monitoring several focal taxa on Barro Colorado Island (BCI) since 2009, including important soil taxa such as ants and termites, as well as some species whose larvae are developing in the soil (Lamarre et al. 2020). A recent study on BCI compared 100 pairwise soil samples either sorted with traditional taxonomy or processed with metabarcoding, considering ants, termites and springtails extracted with Berlese-Tullgren (Bano and Roy 2016) as target taxa (Basset et al. 2020, in prep.). These three taxa represent an important proportion of animal biomass in the soil of tropical rainforests. They also play critical roles in the maintenance and regeneration of the forest, including soil turnover, nutrient cycling and decomposition, plant protection and seed dispersal (Hopkin 1997; Abe et al. 2000; Lach et al. 2010).

The BCI study indicated that (a) a positive correlation existed between the abundance of all species in taxonomic samples and their occurrence in metabarcoding samples. (b) Seasonal shifts in species occurrence and changes in faunal composition between the dry and wet seasons were correlated between taxonomic and metabarcoding samples. (c) False positive and negative species (i.e., species identified positively in the samples but unlikely to be present or species not occurring in the samples but likely to be present) represented a low proportion of species surveyed overall, owing in part to the availability of good reference libraries. These results indicated that metabarcoding could be used for the long-term monitoring of soil arthropods in tropical rainforests (Basset et al. in prep.). However, some challenges emerged. In particular, ant species were not well detected by metabarcoding as compared to samples sorted by traditional taxonomy, as well as compared to the local ant fauna known to inhabit BCI (37% of ant species not detected: Basset et al. in prep.).

It is well-known to entomologists that a range of different methods and traps are needed to provide sound estimates of arthropod diversity in tropical rainforests (Basset et al. 2012). Of interest in this context are the methods, such as Malaise or light traps, able to collect alates of social insects nesting in soil (ants, termites) during dispersal events. Since DNA barcoding can match up different castes within species of social insects (workers, soldiers, alates; Smith et al. 2015), metabarcoding data obtained from Malaise or light traps may also conveniently complement metabarcoding data obtained from soil samples.

The aims of this contribution are twofold. First, we provide a detailed field protocol to collect soil samples and arthropods in a tropical rainforest, to be later processed with metabarcoding. Field protocols to collect soil arthropods and assess soil quality exist (e.g., Römbke et al. 2006), but they are not adapted to processing samples with metabarcoding. Studies targeting the soil fauna and using metabarcoding are typically not adapted to sampling in tropical rainforests, do not provide recommendations for spatial replication, or provide insufficient details about actual sampling and sample handling in the

field (e.g., Arribas et al. 2016; Saitoh et al. 2016; Oliverio et al. 2018; Zinger et al. 2019). In particular, Arribas et al. (2016) advocated using a flotation method, followed by extraction with Berlese-Tullgren apparatus. It is well known that flotation yields low capture rates, is biased against macrofaunal specimens that are less likely to survive the flotation procedure (Macfayden 1953; Southwood and Henderson 2000) and is rather tedious as it involves several filtration procedures (Arribas et al. 2016). In addition, the volume of soil surveyed by flotation is rather small (5 liters in Arribas et al. 2016 against 200 liters in this study, see methods). We therefore hope that details about our protocol will stimulate and guide further studies initiating long-term monitoring of the soil fauna in the tropics. In this article we do not discuss the laboratory protocols related to metabarcoding.

Second, we explore the detectability of soil/litter ants (Formicidae), springtails (Collembola) and termites (Isoptera) with metabarcoding and high-throughput techniques. Specifically, we take advantage of two surveys, performed on BCI and processed with metabarcoding or high-throughput barcoding, to estimate the proportion of soil/litter ant, springtail and termite species from the local fauna that can be detected by metabarcoding or barcoding samples obtained either from Berlese-Tullgren, Malaise or light trap. We then estimate whether each method provides a complementary survey of the local BCI fauna, with respect to the species richness and faunal composition of samples processed by metabarcoding. Eventually, we discuss whether these data can help us to recommend a field protocol with metabarcoding that can detect an adequate number of soil/litter ant, springtail or termite species in this tropical rainforest. In this context, we answer the question "can the soil protocol with Berlese-Tullgren extraction be significantly improved regarding the detectability of focal species by appending one or two additional field protocols, such as Malaise or light traps?"

Material and methods

Study site

Our study was performed on Barro Colorado Island (BCI; 9.15°N, 79.85°W; 120–160 m asl) in Panama. BCI receives an average annual rainfall of 2,662 mm, with an annual average daily maximum and minimum air temperatures of 31.0 °C and 23.6 °C, respectively (http://biogeodb.stri.si.edu/physical_monitoring/research/barro-colorado). The 1,542 ha Barro Colorado Island is covered with lowland tropical forest and was created around 1910, when the Chagres River was dammed to fill the Panama Canal. All samples were obtained from the 50 ha ForestGEO vegetation dynamics plot (or nearby), which is described in Anderson-Teixeira et al. (2015). We surveyed ten locations (500 m sections of trails) inside or near the plot that are used for long-term arthropod monitoring as described in Basset et al. (2013, 2020: Fig. 1, Suppl. material 1: Table S1).

Soil samples

The protocol for obtaining soil samples that were processed with metabarcoding is detailed in Suppl. material 1: Appendix S1 (also published at protocols.io, https://doi.org/10.17504/protocols.io.bj9gkr3w). We provide a summary here, which also included comparing samples processed by traditional taxonomy and samples processed with metabarcoding. Each location was divided into ten sub-locations. For each of the ten locations, we randomly selected five sub-locations. At each sub-location, we took two paired samples 10 cm distant from each other. From this random sampling, we obtained 50 paired soil samples in March 2017, during the dry season. We repeated this sampling protocol in December 2017, during the wet season, and obtained 100 paired soil samples for the two seasons. Paired samples were 50–450 m from each other and were distributed over an area of ca. 60 ha. Each sample consisted of a scoop of soil and litter calibrated to 2 liters, from which the fauna was then extracted with Berlese-Tullgren apparatus (André et al. 2002; Bano and Roy 2016) for 72 hours (see details in Suppl. material 1: Appendix S1). Details of geographic coordinates and other characteristics of the samples may be consulted in Basset et al. (2020: Suppl. material 1: Table S1).

Each pair of samples consisted of two categories: "taxonomic samples", from which the soil fauna was extracted and sorted manually according to morphology and "metabarcoding samples" which were analyzed using DNA metabarcoding. Ants, springtails and termites were identified via morphological and molecular data, as detailed in Basset et al. (2020). Briefly, this included (1) the use of reference collections from the ForestGEO Arthropod Initiative at the Smithsonian Tropical Research Institute (ants and soldier termites) and from the Laboratorio de Ecología y Sistemática de Microartrópodos at the Universidad Nacional Autónoma de México (UNAM, Mexico; springtails); (2) expert opinion; and (3) sequencing the standard DNA barcode region of the gene cytochrome c oxidase subunit I (COI) for a subset of the specimens collected from taxonomic samples. DNA barcoding using Sanger sequencing was conducted at the Centre for Biodiversity Genomics, University of Guelph, using methods described in deWaard et al. (2019a; http://ccdb.ca/resources/). When possible, we sequenced a maximum of five individuals per species or morphospecies (hereafter species). Molecular data were used to confirm identifications based on morphology. Each species was attributed a Barcode Index Number (BIN) according to the Barcode of Life Data System (BOLD; http://www.barcodinglife.org/index.php), which can be used as a proxy taxonomic unit in absence of binomial identification (Ratnasingham and Hebert 2013). In total, we obtained 324 sequences of 171 ants, 114 springtails and 43 termites, which were deposited in existing projects BCIFO, BCICL and BCIIS of BOLD, respectively. A complete list of species of ants, springtails and termites and their BINs recorded in taxonomic samples is detailed in Basset et al. (2020), as well as a discussion about the spatial distribution of species.

Half of the 200 paired samples were processed with metabarcoding at the Biodiversity Institute of Ontario (see laboratory protocols, below).

Malaise trap samples

As part of the Global Malaise trap program (Perez et al. 2017; https://biodiversitygenomics.net/projects/gmp/), a single Malaise trap (Townes 1972) was located near Drayton trail and run weekly for one year from 17 May 2014 to 2 May 2015 (coordinates 09°09'035"N, 79°50'49.6"W, elevation 162 m). Weekly samples were preserved in 95% ethanol and stored at -20 °C. At the end of the collecting period, all bulk samples were shipped to the Centre for Biodiversity Genomics for DNA barcoding. Samples were processed through a high-throughput single molecule, real-time (SMRT) sequencing pipeline implemented on the SEQUEL platform (Hebert et al. 2018). Analysis utilized the same BIN system as indicated above. Out of the 52 weeks (samples) collected, only samples from alternate weeks (n=26) were analyzed. Further, 17 out of these 26 samples included large quantities of springtails (i.e., 1.25 to 41.4 g wet weight). Only subsets of these 17 samples were analyzed for springtails. All Collembola data were uploaded to BOLD and are accessible in the public dataset DS-COLLMAL (650 specimens, 608 sequences, 28 BINs). Specimens were deposited at the Centre for Biodiversity Genomics. Detailed protocols and examples of studies with the Global Malaise Program may be consulted in, e.g., Geiger et al. (2016), Ashfaq et al. (2018) and deWaard et al. (2019b).

Light trap samples

Ten light traps were emplaced in the middle of each trail section where we obtained the previous soil samples. These 10W black light traps of the bucket-type model are described in Lucas et al. (2016). Traps were switched on during the new moon with a timer at 18:00 hours and ran all night long until ca 06:00 hours. One survey consisted of running a trap at each of the ten locations for two non-consecutive nights, yielding 20 night-samples. Over the course of a year, we performed four surveys in March, May, September and December, yielding a total of 80 night-samples. These traps were run during the period 2009–2019. Target taxa were sorted and deposited in the collections of the ForestGEO Arthropod Initiative.

In May 2019, we concurrently ran an additional 10 similar traps that were modified to collect arthropod material in 95% ethanol during two non-consecutive nights. In this case we used laboratory gloves and disinfected all traps tools and recipients with commercial bleach (Clorox de Centroamérica; hypoclorite of sodium 3.5%, hydroxide of sodium 0.3%), after which sampling gear was rinsed with distillated water, to clean bleach residues. Each of the 20 samples obtained was reduced to a volume of 100 ml by plucking one leg of each specimen > 1.5 cm, returning the leg to the sample and discarding the rest of the body. Whole bodies of smaller arthropods were left untouched. Fresh 95% ethanol was added, and samples were stored at -20 °C until shipped to the Biodiversity Institute of Ontario for metabarcoding (see laboratory protocols, below).

Other protocols

As part of the ForestGEO Arthropod Initiative, we used other protocols during 2009–2019 to collect specimens of workers and alates of ants and termites, which were identified and deposited in the collections of the ForestGEO Arthropod Initiative. Parts of these specimens were sequenced at the Biodiversity Institute of Ontario and sequences deposited in BOLD projects BCIFO and BCIIS. Briefly, these protocols included (1) Winkler samples targeting ant workers (Agosti et al. 2000); (2) termite transects targeting termite workers and soldiers (Roisin and Leponce 2004); (3) light traps targeting alates of ants and termites, as described previously; and (4) Malaise traps targeting alates of ants and termites. In the latter case, 10 traps similar to the one described before were set in the South-East corner of the BCI ForestGEO plot. Traps were located at least 200 meters from each other (Barrios and Lagos 2016). They were surveyed weekly from 2002 to 2017, but alates were only sorted from samples covering two weeks of the dry and wet season of each year. In brief, reference collections and DNA barcode libraries were good for ants and termites, owing to the material accumulated over the years and early taxonomical knowledge (e.g., Wheeler 1925; Snyder 1926), whereas springtails were only collected and sequenced from soil samples obtained in this study.

Laboratory protocols and bioinformatics

Detailed laboratory procedures for the soil samples, performed at the Hajibabaei laboratory at the University of Guelph, are indicated elsewhere (Basset et al. in prep.). Briefly, this included extraction of genomic DNA using a DNeasy PowerSoil Kit (Qiagen: Toronto, ON, Canada) according to protocol, and amplification of isolated DNA through a two-stage PCR (Polymerase Chain Reaction) for two amplicons from the DNA barcode region of the COI gene, BR5 and F230R (Folmer 1994; Hajibabaei et al. 2012; Gibson et al. 2014, 2015). The a DNeasy PowerSoil Kit was used because after extraction from Berlese, arthropod specimens were still mixed with small quantities of soil. A second round of PCR with primers was run under the same conditions, using the purified product from the first round of PCR as template. Purified second round PCR product was sequenced on an Illumina MiSeq using the v3 MiSeq sequencing kit. The resulting sequence reads were uploaded into the project MBR-BCI-SOIL of the online platform mBrave ("Multiplex Barcode Research and Visualization Environment", http://www.mbrave.net/; Ratnasingham 2019), for analyzing metabarcoding data and taxonomic assignment. PCR failed for 5 soil samples, hence we report only results for 95 samples. We considered the reference libraries and applied analytical parameters as indicated in Suppl. mate-

rial 1: Table S1. Reference libraries included a mix of mBrave system reference libraries, along with tailored datasets obtained from surveys of the ForestGEO Arthropod Initiative on BCI (Suppl. material 1: Table S1).

Laboratory procedures for samples resulting from the Global Malaise Program were performed at the Canadian Centre for DNA Barcoding and are detailed in Ashfaq et al. (2018), Hebert et al. (2018) and deWaard et al. (2019b). Briefly this included PCR amplification with primers C_LepFolF and C_LepFolR (http://ccdb.ca/resources/), sequencing with an ABI 3730XL and analyses with the BOLD platform (Suppl. material 1: Table S1). For light trap samples, laboratory analyses were similar to those for soil samples, with the following differences. The DNA extracts were used to amplify a 462 bp fragment of the COI barcode region using insect primers AncientLepF3/C_LepFolR. Sequencing was performed on an Ion Torrent S5 high-throughput sequencer. The resulting sequence reads were analyzed using the platform mBrave with analytical parameters and reference libraries as indicated in Suppl. material 1: Table S1. Thus, for an objective comparison among sampling methods, we considered for all samples the detected BINs of species or morphospecies from our reference libraries. We do not discuss here the unknown species that currently lack BINs, but which may also have been detected by metabarcoding.

Statistical analyses

For the purpose of long-term monitoring, common species (as opposed to rare species) are most likely to be amenable to statistical analysis (Basset et al. 2013). Hence, we analyzed all the species but focused our results on the most common species. They were defined as being detected in at least 10 samples for one of the three collecting methods.

In other terms, for a particular method and species, the probability to detect at least one specimen at each of the locations should be p=1.0.

Information about nesting sites and the commonness of each species are indicated in Suppl. material 2: Appendix S2.

To visualize the number of species of ants, springtails and termites detected with the three collecting methods, we used area-proportional Venn diagrams (i.e., Euler diagrams) drawn with the R package 'eulerr' (Larsson 2020). We drew these diagrams (i) for all species, (ii) for categories of social insects nesting either in the soil or in arboreal habitats, (iii) as well as for common species nesting (ants, termites) or dwelling (springtails) in the soil. Ants and termites were assigned to soil- or arboreal-nesting categories according to AntWiki (AntWiki 2020), other published literature and our general understanding of the BCI ants. Soil-nesting included the following nesting sites: hypogaeic, epigaeic, under stones, dead wood and litter. When species were only identified by their BIN, we assigned the same nesting categories as to congeneric species with available information.

The data for Malaise and, particularly, light traps are conservative because sample size was unequal (95 soil samples; 182 trap-days for Malaise and 20 trap-nights for light traps) and lower than what would be intended to perform within a year (see below). To reduce the effect of sampling effort on species richness, we compared the three methods with species accumulation and rarefaction curves, separately for each target taxa. We considered the occurrence of species in samples (i.e., the sum of the times a species was detected in all samples available) and computed rarefaction curves of species richness vs. the number of samples with the R package 'iNEXT' (Hsieh et al. 2016). To help visualize results of the rarefaction analyses, we extrapolated the curves of species richness with the same package to 150 samples. We also computed an estimate of total species richness for each method and taxa with iNEXT. We paid attention to the species richness accumulated with each collecting method with reference to the ideal minimum number of samples that should be performed within a year to monitor adequately common species (see discussion in Basset et al. 2013). These values are 100 soil samples (Basset et al. 2020), 80 light trap-night samples (Basset et al. 2017) and estimated to be 40 one-week Malaise trap samples. We also computed species accumulation curves for common species nesting or dwelling in the soil, as defined previously.

To evaluate differences in the faunal composition of samples obtained with the three collecting methods, we performed non-metric multidimensional scaling (NMDS; calculated with Jaccard similarity) with the function 'metaMDS' of the R package 'vegan' (Oksanen et al. 2018). We performed four ordinations separately, for all target species together, for ants, springtails and termites. We considered a species x samples matrix with presence-absence data (252 species × 141 samples for all species together). We used the function 'ordiellipse' of the vegan package to draw ellipses representing 95% CI around the centroids. To test for differences between groups (methods) we used a permutational multivariate analysis of variance using distance matrices, which was performed with the function 'adonis' of vegan (Oksanen et al. 2018).

Results

All species occurrence in soil, Malaise and light trap samples are detailed in Suppl. material 2: Appendix S2. A nearly equal number of species was detected in Malaise trap and soil samples, whereas only half of that number was detected in light trap samples (Table 1). However, the sampling effort with light traps was clearly lower than for other methods (Table 1). Each collecting method detected a rather distinct fauna (Fig. 1). Samples from the Malaise traps detected about twice the number of ant species than the soil and light trap samples (Fig. 1a). A high number of springtail species were detected in soil samples, followed by Malaise samples, but very few species were detected in light trap samples (Fig. 1b). Termites were best detected in soil samples, and alates of termites were better detected in light trap than in Malaise samples (Fig. 1c). Overall, for the three target taxa, 87 species were unique-

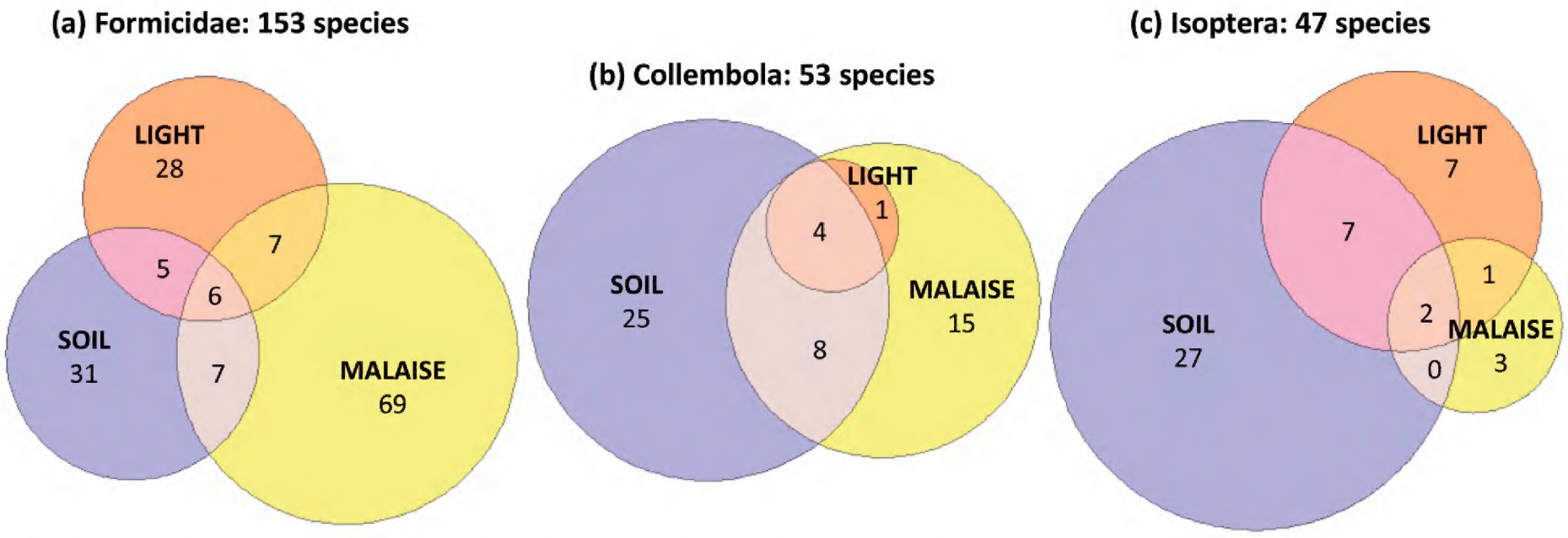

Figure 1. Euler diagrams for (a) ants, (b) springtails and (c) termites indicating the number of species detected in metabarcoding samples by each collecting method.

Table 1. Characteristics of samples obtained with each collecting method.

Variable	Soil samples	Malaise trap	Light trap
Year of collecting	2017	2014	2019
No. samples	100	26	20
No. trap-days	–	182	20
Spatial replication	10 locations	None	10 locations
Seasonal replication	Dry and wet season	Dry and wet season	Wet season only
No. sequences*	3,473,116	1,867	36,996
No. species detected (BINs)*	120	123	68

* Only 3 focal taxa, ants, springtails and termites

ly detected by Malaise samples, 83 by soil samples and 35 by light trap samples (Fig. 1).

Ant species nesting in the soil were similarly well detected in all three types of samples, along the series Malaise trap > soil samples > light trap, whereas arboreal ants were mostly detected in Malaise trap samples (Suppl. material 1: Fig. S1). The pattern was rather different for termites nesting in the soil, which were mostly detected in soil samples. Arboreal termites were mostly detected in light trap samples (Suppl. material 1: Fig. S1). The pattern was also different for common species nesting or dwelling in the soil, which represented 22–36% of all species detected, depending on the taxa (Suppl. material 1: Fig. S2). The most common ant species were detected in light trap and soil samples. Common species of springtails were mostly detected in soil samples, whereas common species of soil termites were well detected in soil samples and light trap samples.

Species accumulated faster in Malaise trap samples for ants (total estimated species richness with iNEXT = 168 ± 32 [s.e.]) than for soil (74 ± 15) and light trap samples (54 ± 7; Fig. 2). This was also apparent when considering the minimum number of samples to be performed

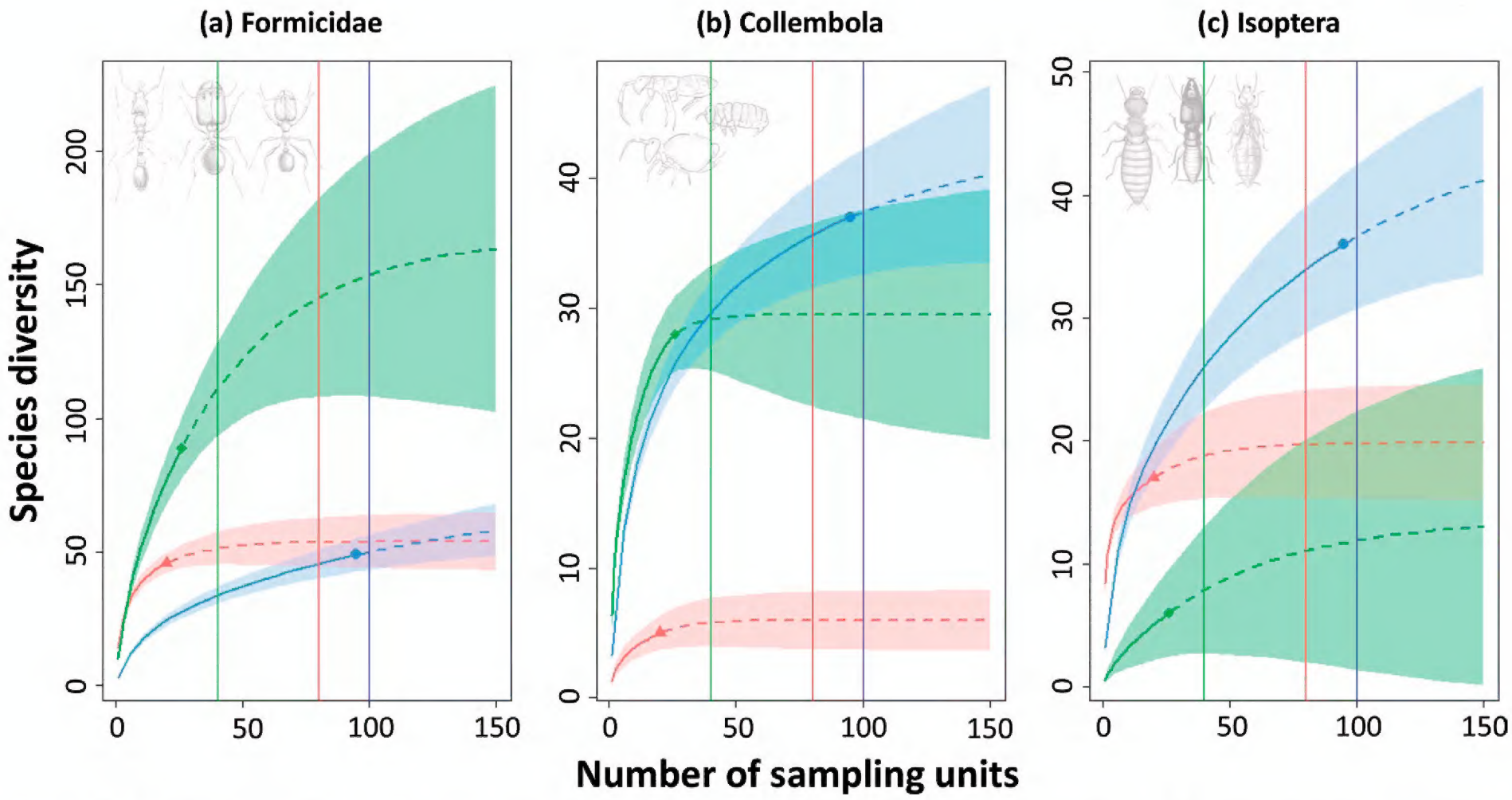

Figure 2. Accumulation of species richness vs. the number of samples for (a) ants, (b) springtails and (c) termites, detailed for each collecting method: blue = soil samples; green = Malaise trap samples; red = light trap samples. Solid lines are interpolated curves, dotted lines are extrapolated curves to 150 samples. The vertical lines indicate, for each collecting method, the targeted minimum number of samples within one year (see text).

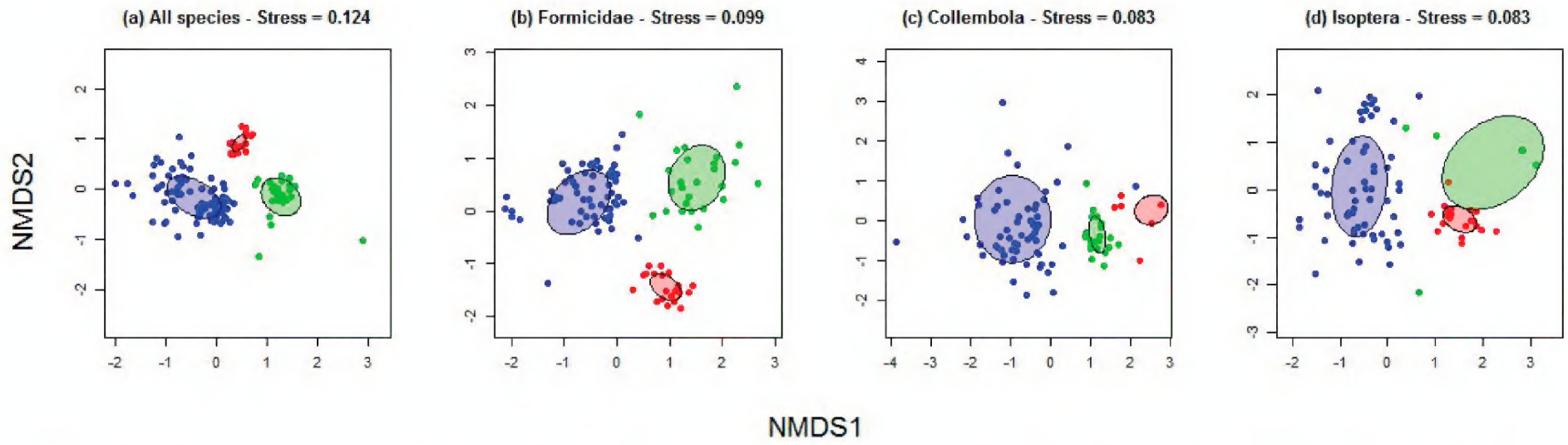

Figure 3. NMDS plots for (a) all species, (b) Formicidae, (c) Collembola and (d) Isoptera. Plots of samples (blue = soil samples, green = Malaise trap samples, red = light trap samples) in the first two axes of the ordinations. The ellipses represent 95% confidence limits around the centroids of each method.

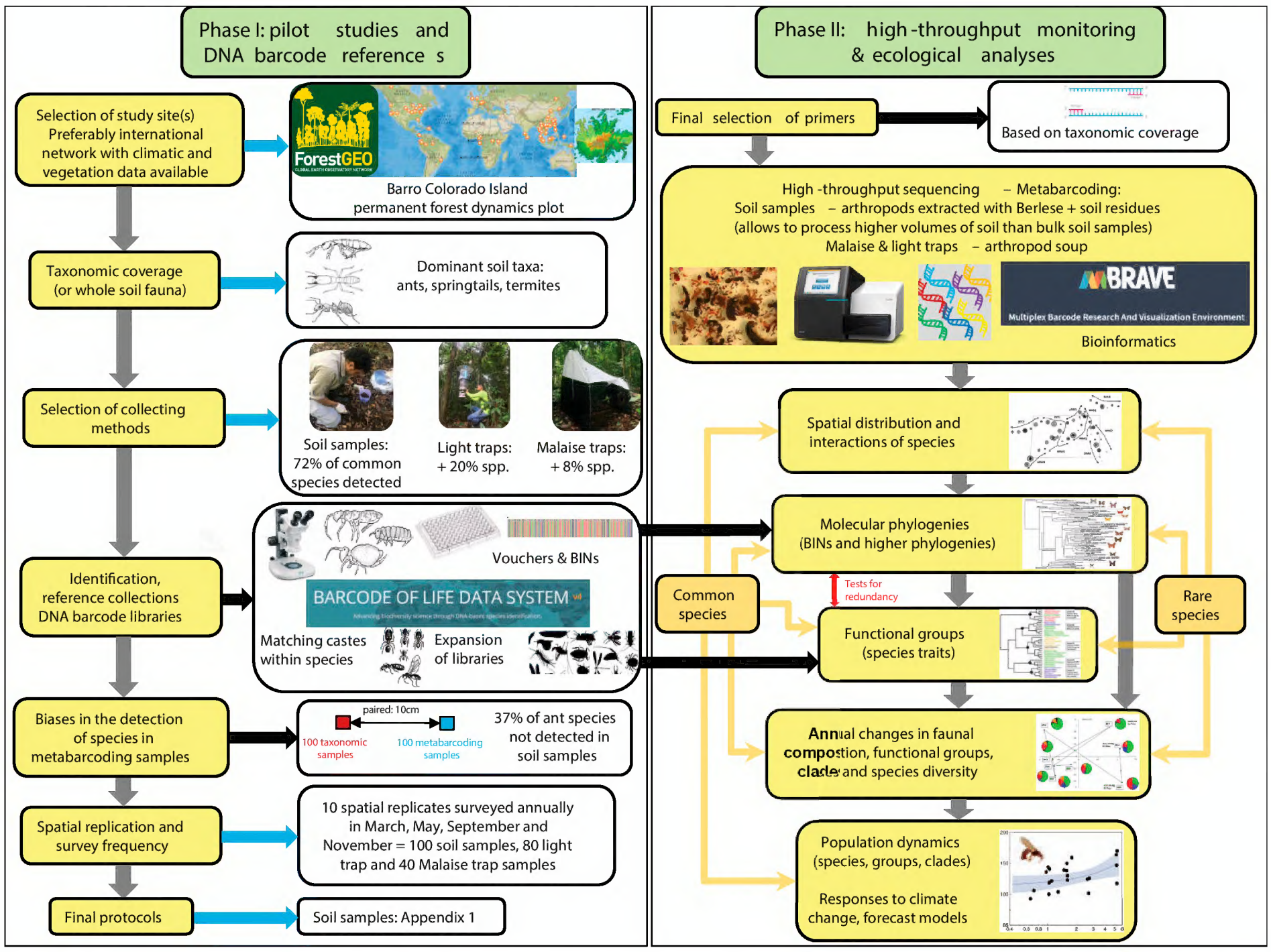
Phase I: pilot studies and DNA barcode reference s
Phase II: high-throughput monitoring & ecological analyses
Selection of study site(s) Preferably international network with climatic and vegetation data available
ForestGEO
Barro Colorado Island permanent forest dynamics plot
Taxonomic coverage (or whole soil fauna)
Dominant soil taxa: ants, springtails, termites
Selection of collecting methods
Soil samples: 72% of common species detected
Light traps: + 20% spp.
Malaise traps: + 8% spp.
Identification, reference collections DNA barcode libraries
Vouchers & BINs
BARCODE OF LIFE DATA SYSTEM
Matching castes within species
Expansion of libraries
Biases in the detection of species in metabarcoding samples
paired: 10cm
100 taxonomic samples
100 metabarcoding samples
37% of ant species not detected in soil samples
Spatial replication and survey frequency
10 spatial replicates surveyed annually in March, May, September and November = 100 soil samples, 80 light trap and 40 Malaise trap samples
Final protocols
Soil samples: Appendix 1
Final selection of primers
Based on taxonomic coverage
High-throughput sequencing – Metabarcoding:
Soil samples – arthropods extracted with Berlese + soil residues
(allows to process higher volumes of soil than bulk soil samples)
Malaise & light traps – arthropod soup
BRAVE
Multiplex Barcode Research And Visualization Environment
Bioinformatics
Spatial distribution and interactions of species
Molecular phylogenies (BINs and higher phylogenies)
Tests for redundancy
Common species
Rare species
Functional groups (species traits)
Annual changes in faunal composition, functional groups, clades and species diversity
Population dynamics (species, groups, clades)
Responses to climate change, forecast models

Figure 4. Long-term monitoring scheme for the soil arthropods in the tropical rainforest of Barro Colorado Island, Panama. Horizontal arrows colored in blue refer specifically to this study, pointing to results).

within one year for each collecting method. By contrast, springtails showed faster species accumulation in soil samples (estimated richness 43 ± 6), followed by Malaise trap samples (30 ± 2) and light trap samples (6 ± 2). The pattern was also different for termites, which showed faster species accumulation in soil samples (estimated richness 48 ± 9), followed by light trap samples (20 ± 4) and Malaise trap samples (14 ± 11; Fig. 2). For common species nesting or dwelling in the soil equivalent species accumulation curves were rather different. For Formicidae, common species accumulated faster in samples in the series light trap > soil samples > Malaise trap (Suppl. material 1: Fig. S3). For springtails the equivalent series was soil samples > Malaise trap > light trap and for termites it was soil samples > light trap (sampling effort was too low to compute an accumulation curve for termites collected by Malaise trap; Suppl. material 1: Fig. S3).

The NMDS plots confirmed that faunal composition was rather different in samples obtained by the three collecting methods. Differences in faunal composition were most marked when all species were considered together ($F_{2,135} = 29.17, p = 0.001$; Fig. 3), then for ants and spring-

tails ($F_{2,123} = 29.74$, $p = 0.001$ and $F_{2,109} = 29.76$, $p = 0.001$, respectively), and for termites ($F_{2,115} = 26.56$, $p = 0.001$).

Discussion

Entomologists are well aware that an array of collecting methods are necessary to survey most arthropod species thriving in even relatively small areas (Basset et al. 2012), and this is also true for soil arthropods (André et al. 2002). These challenges are compounded when processing samples by metabarcoding because detection rates may differ among species (Saitoh et al. 2016; Elbrecht et al. 2019). Hence, the faunal composition of soil samples sorted manually may differ from that in soil samples processed with metabarcoding. This is not necessarily detrimental to long-term monitoring, but the possible bias and limitations of metabarcoding must be acknowledged, and perhaps corrected either in the field, laboratory or via bioinformatics. Here, we are concerned with field protocol and provided a detailed protocol to collect soil samples and arthropods in a tropical rainforest to be processed with metabarcoding. We then evaluated whether samples processed with metabarcoding and obtained with other methods commonly used by entomologists, Malaise and light traps, may be able to detect a substantial number of soil-dwelling species not detected in soil samples, for the target taxa of Formicidae, Collembola and Isoptera in this tropical rainforest. We show that a substantial number of additional species are detected in metabarcoding samples obtained with Malaise and light traps, including species that are common and amenable to statistical analyses of long-term trends. This validates the approach of using different field protocols to obtain samples processed by metabarcoding with the purpose of monitoring in the long term a reasonable share of the focal species inhabiting BCI.

We restricted our study to three dominant taxa of soil/litter arthropods. We do not have sufficient information to discuss cogently how well metabarcoding may detect other invertebrates in the soils of BCI, for lack of sound DNA libraries for these other taxa. Among others they may include earthworms, mites, spiders, myriapods, nematodes and annelids.

We showed that each collecting method detected a rather distinct fauna, in terms of species richness or faunal composition. Overall, when considering the three target taxa, soil and Malaise trap samples detected 213 species (73%) of all species detected by the three collecting methods. However, patterns were different between taxa, with Malaise trap samples detecting many ant species, whereas soil samples were more efficient at detecting springtail and termite species. Both Malaise and light traps detected a high proportion of arboreal ants and termites, representing 35% of the total of species of social insects detected. Further, with respect to long-term monitoring of soil-nesting or soil-dwelling common species (more amenable to statistical trends), the best combination of two methods, in terms of maximizing the detection of common species, were soil and light trap samples, detecting 94% of the total of the species common in the soil. This emphasizes that different methods may be suitable for different research goals. For a snapshot survey of the local area (including the soil and arboreal fauna), metabarcoding samples obtained with all three collecting methods would obviously be best, followed by a combination of soil plus Malaise samples, if funding is limited. For generating time-series in the long-term, soil samples detected 72% of common species and would be the choice method, followed by a combination of soil plus light trap samples. These results were not too different when considering species accumulated and the ideal minimum number of samples that should be performed within a year to monitor common species adequately (Basset et al. 2013).

Several factors may hinder a direct comparison of the data available, even when considering the number of samples collected in species accumulation curves. First, we cannot consider reliable estimates of species abundance in samples processed with metabarcoding with the present technology (Lamb et al. 2018; Creedy et al. 2019). Second, it is difficult to compare samples from the different methods with respect to spatial (soil samples > light trap > Malaise trap and seasonal replication (Malaise trap > soil samples > light trap). Hence, we can only conjecture what would happen in our BCI system if we were to obtain with an ideal protocol each year 100 soil samples (10 spatial replicates, two sampling seasons), 40 one-week Malaise trap samples (10 replicates, 4 annual surveys) and 80 light trap-night samples (10 two-nights replicates, 4 annual surveys). In this case, we can assume that (1) more ant species nesting in the soil would be detected in our samples (the soil ant species detected in this study represent about 50% of the total species richness of soil-nesting ants on BCI; D. Donoso, unpubl. data); (2) we may also detect more springtail species with more replicates of Malaise trap (and complete analysis of springtails, see methods), although not all these species may be dwelling in the soil/litter (see below); and (3) a higher number of species may be amenable to statistical analysis (common species). Third, de Kerdrel et al. (2020) noted, in their comparison of arthropod samples obtained with different methods and processed with high-throughput barcoding, that fewer sequences were recovered from soil trapping methods (Berlese and pitfall traps) as compared to methods targeting flying insects or arthropods perched on vegetation (Malaise traps, beating). These authors suggested that samples with a considerable amount of soil and debris may include compounds that may inhibit PCR. Water in the soil may also increase DNA degradation (de Kerdrel et al. 2020).

We used high-throughput barcoding for Malaise trap samples instead of the metabarcoding used for soil and light trap samples. Because longer sequences were obtained with the former, BIN assignment was probably more accurate for Malaise trap samples, resulting in better quality data (deWaard et al. 2019a). Further, in this study, we compared species with BINs among methods for a convenient and straightforward approach, as op-

posed to also considering operational taxonomic units (OTUs) of unknown species without BINs. It is legitimate to also ask whether our results would be different when considering these possible extra OTUs. We believe that at least for ants and termites, differences would have been negligible because of the near complete DNA reference libraries available for these taxa inhabiting BCI.

Irrespective of the collecting method used, the number of ant species recovered by metabarcoding remained rather low. This may be related to lower extraction rate and sequence recovery. For instance, out of 14 arthropod orders screened by de Kerdrel et al. (2020), Hymenoptera (including ants) showed the lowest sequence recovery rate. Implementing our three collecting methods with an ideal sample size of 100 soil, 40 Malaise and 80 light trap samples annually, would generate reliable time-series for the 24 ant species that we labeled as being common and amenable to statistical analysis of long-term trends. These 24 species represent 11% of the total number of soil/litter ant species inhabiting BCI. This proportion is similar or higher to that for the soil/litter ant species amenable to similar statistical analysis (8%) when acquiring each year 50 Winkler samples sorted manually. Replacing Winkler samples by soil samples and adding Malaise trap samples would increase field work slightly (i.e., human labor to survey traps and handle samples) as compared to our current annual sampling scheme (Lamarre et al. 2020). However, this would decrease the time related to specimen processing and identification considerably, not to mention the value added of obtaining metabarcoding data of non-target taxa for all three collecting methods.

In this contribution, springtails were all assumed to be soil-dwelling. This is certainly an oversimplification, given that many species also thrive in arboreal habitats, even though their biology remains poorly known (Hopkin 1997). Although springtails are often neglected in Malaise trap samples, they are by no means infrequent in these samples (Greenslade and Florentine 2013) and can locally represent a substantial proportion of catches albeit with low species richness (Palacios-Vargas and Gómez-Anaya 1993; Palacios-Vargas et al. 1998). This concerns mainly the family Entomobryidae and in the Neotropics genera such as *Seira*, *Entomobrya* or *Lepidocyrtus*, as well as Paronellidae such as *Trogolaphysa* and often *Salina*, which may climb on the mesh of the trap (J.G. Palacios-Vargas, pers. obs.). By contrast, this is hardly possible in light traps as they were suspended at 1.3 m height (Lucas et al. 2016). While both springtails and termites may fall victim to predatory ants (Basset et al. 2020), and thus be detected in metabarcoding samples, this is not necessarily detrimental to monitoring, as it ascertains the presence of the prey in samples. There are at least 104 species of springtails on BCI (Basset et al. 2020), but currently only 33% of these species have BINs. Hence, we can conservatively estimate that a sampling scheme based on 100 soil samples annually would allow for the monitoring of at least the 19 species that we labeled as being common and amenable to statistical analysis of long-term trends (18% of the total number of species). This proportion can probably be increased as springtail BINs are added to the reference libraries.

There are at least 62 species of termites on BCI (Basset et al. 2020; Y, Basset et al., unpubl. data), including about 31% of specialized humus-feeding species (mostly Apicotermini) species, 53% of wood-eating species in the understorey and 16% of wood-eating arboreal species. Soil metabarcoding samples detected many of the humus-feeding species. A protocol based on 100 soil and 80 light trap samples annually would probably allow for the monitoring of about 16% of termite species (the 10 soil species labeled as common), although wood-eating species may be underrepresented. Light trap data would allow for monitoring alates of a few more arboreal species.

When time and funding do not allow the combined use of the three collecting methods tested in this study, Malaise traps have the advantage over light traps of not requiring electrical power in the field, as well as requiring less time to set up and run in the field. Yang et al. (2014) proposed that soil and Malaise trap samples processed with metabarcoding be tested as a candidate method for rapid environmental monitoring in terrestrial ecosystems. Relevant to this, it is also legitimate to ask whether it is sound to monitor workers (mostly in soil samples) or alates (mostly Malaise and light trap samples) of social insects, since metabarcoding does not allow for distinguishing between the two casts (but barcoding does after verification of vouchers). For example, out of the +16,000 specimens of ants sorted from the 10 BCI Malaise traps surveyed during the period of 2002–2017, 95.54% were male alates, 4.40% were female alates and only 0.06% were workers (D. Donoso et al. unpubl. data). There are undeniable advantages to monitor alates as, in some cases, workers are far more difficult to survey than alates. Because workers' swarms are constantly on the move, the spatial distribution of army ants is highly patchy (Kaspari et al. 2011) and they are difficult to survey in soil or Winkler samples, whereas alates are readily caught in numbers in light traps. Likewise, all termites collected in light traps samples are alates and include a good proportion (60%) of species that are arboreal and hence rarely collected in monitoring samples obtained from the forest understorey.

Conclusion

Based on our results, a multiple protocol including 100 soil, 40 Malaise and 80 light trap samples processed annually by metabarcoding would allow monitoring of at least 11%, 18% and 16% of species of soil/litter ants, springtails and termites, respectively, present on BCI. All these species have been formally vouchered, and many are identified (Suppl. material 2: Appendix S2), so that in the future it may be possible to discuss cogently potential reasons for population decline or increase over time. Although these common species may not represent a large share of species richness, they represent a high proportion

of the BCI fauna in terms of abundance. For example, common soil/litter ant species as defined in this study represent about 80% of all ant individuals collected annually by Winkler samples for our long-term monitoring scheme. The sampling effort suggested may be suitable for the area of BCI but may vary for other rainforest locations. Figure 4 summarizes how we envision the long-term monitoring of soil arthropods at BCI. This provides a starting point for pilot studies concerned with the long-term monitoring of arthropods in tropical rainforests.

Acknowledgments

We thank ForestGEO and the Smithsonian Tropical Research Institute for logistical support. Richard Ctvrtecka, Petr Blazek and Pamela Polanco helped with light trap samples. Scott Miller commented on an early draft of the manuscript.

This study was supported by SENACYT (FID16-070 to YB and HB), the Czech Science Foundation (GAČR 20-31295S to YB) and ForestGEO. Grants from the Smithsonian Institution Barcoding Opportunity FY012, FY015 and FY018 (to YB) and in-kind help from the Canadian Centre for DNA Barcoding as part of the iBOL program allowed sequencing specimens. YB and HB were supported by the Sistema Nacional de Investigación, SENACYT, Panama. G.P.A.L. was supported by a GAČR Junior Grant (19-15645Y).

The field protocol was deposited at protocols.io (dx.doi.org/10.17504/protocols.io.bj9gkr3w). Sequences and BINs used as reference libraries are publicly accessible in BOLD projects as indicated in the text, the dataset for Collembola as dx.doi.org/10.5883/DS-COLLMAL. Sequences are also deposited in GenBank databases under accession numbers KP845288–KP849461, KU745531–KU745532, KX072335–KX072563, MF922335–MF970719, MG03-0727, MK758129–MK770080, MN345316–MN621065 and MT357731. Suppl. material 2: Appendix S2 provides all the information about the identity of species and their occurrence in samples. The raw sequences were deposited at the Sequence Read Archive (SRA; https://www.ncbi.nlm.nih.gov/bioproject/PRJNA668155).

References

Abe T, Bignell DE, Higashi M [Eds] (2000) Termites: Evolution, Sociality, Symbiosis, Ecology, Kluwer Academic Publishers, Dordrecht. https://doi.org/10.1007/978-94-017-3223-9

Agosti D, Majer JD, Alonso LE, Schultz TR (2000) Ants: Standard Methods for Measuring and Monitoring Biodiversity. Smithsonian Institution Press, Washington DC.

Anderson-Teixeira KJ, Davies SJ, Bennett AC, Gonzalez-Akre EB, Muller-Landau HC, Wright SJ, Salim KA, Zambrano AMA, Alonso A, Baltzer JL, Basset Y, Bourg NA, Broadbent EN, Brockelman WY, Bunyavejchewin S, Burslem DFRP, Butt N, Cao M, Cardenas D, Chuyong GB, Clay K, Cordell S, Dattaraja HS, Deng X, Detto M, Du X, Duque A, Erikson DL, Ewango CEN, Fischer GA, Fletcher C, Foster RB, Giardina CP, Gilbert GS, Gunatilleke N, Gunatilleke S, Hao Z, Hargrove WW, Hart TB, Hau BCH, He F, Hoffman FM, Howe RW, Hubbell SP, Inman-Narahari FM, Jansen PA, Jiang M, Johnson DJ, Kanzaki M, Kassim AR, Kenfack D, Kibet S, Kinnaird MF, Korte L, Kral K, Kumar J, Larson AJ, Li Y, Li X, Liu S, Lum SKY, Lutz JA, Ma K, Maddalena DM, Makana J-R, Malhi Y, Marthews T, Serudin RM, McMahon SM, McShea WJ, Memiaghe HR, Mi X, Mizuno T, Morecroft M, Myers JA, Novotny V, de Oliveira AA, Ong PS, Orwig DA, Ostertag R, den Ouden J, Parker GG, Phillips RP, Sack L, Sainge MN, Sang W, Sri-ngernyuang K, Sukumar R, Sun I-F, Sungpalee W, Suresh HS, Tan S, Thomas SC, Thomas DW, Thompson J, Turner BL, Uriarte M, Valencia R, Vallejo MI, Vicentini A, Vrška T, Wang X, Wang X, Weiblen G, Wolf A, Xu H, Yap S, Zimmerman J (2015) CTFS-Forest GEO: a worldwide network monitoring forests in an era of global change. Global Change Biology 21: 528–549. https://doi.org/10.1111/gcb.12712

André HM, Ducarme X, Lebrun P (2002) Soil biodiversity: myth, reality or conning? Oikos 96: 3–24. https://doi.org/10.1034/j.1600-0706.2002.11216.x

AntWiki (2020) AntWiki https://www.antwiki.org/wiki/Welcome_to_AntWiki [retrieved on 10 August 2020]

Arribas P, Andújar C, Hopkins K, Shepherd M, Vogler AP (2016) Metabarcoding and mitochondrial metagenomics of endogean arthropods to unveil the mesofauna of the soil. Methods in Ecology and Evolution 7: 1071–1081. https://doi.org/10.1111/2041-210X.12557

Ashfaq M, Sabir JSM, El-Ansary HO, Perez K, Levesque-Beaudin V, Khan AM, Rasool A, Gallant C, Addesi J, Hebert PDN (2018) Insect diversity in the Saharo-Arabian region: Revealing a little-studied fauna by DNA barcoding. PLoS ONE 13: e0199965. https://doi.org/10.1371/journal.pone.0199965

Bano R, Roy S (2016) Extraction of Soil Microarthropods: A low cost Berlese-Tullgren funnels extractor. International Journal of Fauna and Biological Studies 2: 14–17. http://www.faunajournal.com/archives/?year=2016&vol=3&issue=2&part=A&ArticleId=193

Barrios H, Lagos M (2016) Cambios en la estructura de la comunidad de Cerambycidae (Coleoptera) en la Isla de Barro Colorado, Panamá. Scientia 26: 7–24.

Basset Y, Cizek L, Cuénoud P, Didham RK, Guilhaumon F, Missa O, Novotny V, Ødegaard F, Roslin T, Schmidl J, Tishechkin AK, Winchester NN, Roubik DW, Aberlenc H-P, Bail J, Barrios H, Bridle JR, Castaño-Meneses G, Corbara B, Curletti G, da Rocha WD, De Bakker D, Delabie JHC, Dejean A, Fagan LL, Floren A, Kitching RL, Medianero E, Miller SE, de Oliveira EG, Orivel J, Pollet M, Rapp M, Ribeiro SP, Roisin Y, Schmidt JB, Sørensen L, Leponce M (2012) Arthropod diversity in a tropical forest. Science 338: 1481–1484. https://doi.org/10.1126/science.1226727

Basset Y, Eastwood R, Sam L, Lohman DJ, Novotny V, Treuer T, Miller SE, Weiblen GD, Pierce NE, Bunyavejchewin S, Sakchoowong W, Kongnoo P, Osorio-Arenas MA (2013) Cross-continental comparisons of butterfly assemblages in rainforests: implications for biological monitoring. Insect Conservation and Diversity 6: 223–233. https://doi.org/10.1111/j.1752-4598.2012.00205.x

Basset Y, Lamarre GPA (2019) Toward a world that values insects. Science 364: 1230–1231. https://doi.org/10.1126/science.aaw7071

Basset Y, Lamarre GPA, Ratz T, Segar ST, Decaëns T, Rougerie R, Miller SE, Perez F, Bobadilla R, Lopez Y, Ramirez JA, Aiello A, Barrios H (2017) The Saturniidae of Barro Colorado Island, Panama: A mod-

el taxon for studying the long-term effects of climate change? Ecology and Evolution 7: 9991–10004. https://doi.org/10.1002/ece3.3515

Basset Y, Miller SE, Gripenberg S, Ctvrtecka R, Dahl C, Leather SR, Didham RK (2019) An entomocentric view of the Janzen-Connell hypothesis. Insect Conservation and Diversity 12: 1–8. https://doi.org/10.1111/icad.12337

Basset Y, Palacios-Vargas JG, Donoso DA, Castaño-Meneses G, Decaëns T, Lamarre GPA, De León LF, Rivera M, García-Gómez A, Perez F, Bobadilla R, Lopez Y, Ramirez JA, Cruz MM, Galván AA, Mejía-Recamiere BE, Barrios H (2020) Enemy-free space and the distribution of ants, springtails and termites in the soil of one tropical rainforest. European Journal of Soil Biology 99: 103193. https://doi.org/10.1016/j.ejsobi.2020.103193

Beng KC, Tomlinson KW, Shen XH, Surget-Groba Y, Hughes AC, Corlett RT, Slik JF (2016) The utility of DNA metabarcoding for studying the response of arthropod diversity and composition to land-use change in the tropics. Scientific Reports 6: 1–13. https://doi.org/10.1038/srep24965

Creedy TJ, Ng WS, Vogler AP (2019) Toward accurate species-level metabarcoding of arthropod communities from the tropical forest canopy. Ecology and Evolution 9: 3105–3116. https://doi.org/10.1002/ece3.4839

Crossley MS, Meier AR, Baldwin EM, Berry LL, Crenshaw LC, Hartman GL, Lagos-Kutz D, Nichols DH, Patel K, Varriano S, Snyder WE, Moran MD (2020) No net insect abundance and diversity declines across US Long Term Ecological Research sites. Nature Ecology & Evolution 4: 1368–1376. https://doi.org/10.1038/s41559-020-1269-4

Deutsch CA, Tewksbury JJ, Huey RB, Sheldon KS, Ghalambor CK, Haak DC, Martin PR (2008) Impacts of climate warming on terrestrial ectotherms across latitude. Proceedings of the National Academy of Sciences 105: 6668–6672. https://doi.org/10.1073/pnas.0709472105

de Kerdrel GA, Andersen JC, Kennedy SR, Gillespie R, Henrik K (2020) Rapid and cost-effective generation of single specimen multilocus barcoding data from whole arthropod communities by multiple levels of multiplexing. Scientific Reports 10: 1–78. https://doi.org/10.1038/s41598-019-54927-z

deWaard JR, Levesque-Beaudin V, deWaard SL, Ivanova NV, McKeown JT, Miskie R, Naik S, Perez KHJ, Ratnasingham S, Sobel CN, Sones JE, Steinke C, Telfer AC, Young AD, Young MR, Zakharov EV, Hebert PDN (2019a) Expedited assessment of terrestrial arthropod diversity by coupling Malaise traps with DNA barcoding. Genome 62: 85–95. https://doi.org/10.1139/gen-2018-0093

deWaard JR, Ratnasingham S, Zakharov EV, Borisenko AV, Steinke D, Telfer AC, Perez KHJ, Sones JE, Young MR, Levesque-Beaudin V, Sobel CN, Abrahamyan A, Bessonov K, Blagoev G, deWaard SL, Ho C, Ivanova NV, Layton KKS, Lu L, Manjunath R, McKeown JTA, Milton MA, Miskie R, Monkhouse N, Naik S, Nikolova N, Pentinsaari M, Prosser SWJ, Radulovici AE, Steinke C, Warne CP, Hebert PDN (2019b) A reference library for Canadian invertebrates with 1.5 million barcodes, voucher specimens, and DNA samples. Scientific Data 6: 308. https://doi.org/10.1038/s41597-019-0320-2

Didham RK, Basset Y, Collins CM, Leather SR, Littlewood NA, Menz MHM, Müller J, Packer L, Saunders ME, Schönrogge K, Stewart AJA, Yanoviak SP, Hassall C (2020) Interpreting insect declines: seven challenges and a way forward. Insect Conservation and Diversity 13: 103–114. https://doi.org/10.1111/icad.12408

Donoso DA (2017) Tropical ant communities are in long-term equilibrium. Ecological Indicators 83: 515–523. https://doi.org/10.1016/j.ecolind.2017.03.022

Elbrecht V, Braukmann TW, Ivanova NV, Prosser SW, Hajibabaei M, Wright M, Zakharov EV, Hebert PDN, Steinke D (2019) Validation of COI metabarcoding primers for terrestrial arthropods. PeerJ 7: e7745. https://doi.org/10.7717/peerj.7745

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3: 294–299.

Geiger M, Moriniere J, Hausmann A, Haszprunar G, Wägele W, Hebert P, Rulik B (2016) Testing the Global Malaise Trap Program – How well does the current barcode reference library identify flying insects in Germany? Biodiversity Data Journal 4: e10671. https://doi.org/10.3897/BDJ.4.e10671

Gibson JF, Shokralla S, Curry C, Baird DJ, Monk WA, King I, Hajibabaei M (2015) Large-scale biomonitoring of remote and threatened ecosystems via high-throughput sequencing. PLoS ONE 10: e0138432. https://doi.org/10.1371/journal.pone.0138432

Gibson J, Shokralla S, Porter TM, King I, van Konynenburg S, Janzen DH, Hallwachs W, Hajibabaei M (2014) Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through DNA metasystematics. Proceedings of the National Academy of Sciences 111: 8007–8012. https://doi.org/10.1073/pnas.1406468111

Goulson D, Nicholls E, Botías C, Rotheray EL (2015) Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. Science 347: e1255957. https://doi.org/10.1126/science.1255957

Greenslade P, Florentine S (2013) 6 COLLEMBOLA (springtails). In: Driessen MM, Mallick SA (Eds) The distributions of invertebrate species along the Warra-Mount Weld Altitudinal Transect in 2001–2002 and identification of taxa restricted by altitude. Nature Conservation Report 13/4, Department of Primary Industries, Parks, Water and Environment, Hobart, 24–38.

Hallmann CA, Sorg M, Jongejans E, Siepel H, Hofland N, Schwan H, Stenmans W, Müller A, Sumser H, Hörren T, Goulson D, de Kroon H (2017) More than 75 percent decline over 27 years in total flying insect biomass in protected areas. PLoS ONE 12: e0185809. https://doi.org/10.1371/journal.pone.0185809

Hajibabaei M, Spall JL, Shokralla S, van Konynenburg S (2012) Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol. BMC Ecology 12: 1–28. https://doi.org/10.1186/1472-6785-12-28

Hebert PD, Braukmann TW, Prosser SW, Ratnasingham S, DeWaard JR, Ivanova V, Janzen DH, Hallwachs W, Naik S, Sones JE, Zakharov EV (2018) A Sequel to Sanger: amplicon sequencing that scales. BMC genomics 19: 1–14. https://doi.org/10.1186/s12864-018-4611-3

Hopkin SP (1997) The Biology of Springtails (Insects: Collembolan). Oxford University Press, Oxford.

Hsieh TC, Ma KH, Chao A (2016) iNEXT: An R package for rarefaction and extrapolation of species diversity (Hill numbers). Methods in Ecology and Evolution 7: 1451–1456. https://doi.org/10.1111/2041-210X.12613

Kaspari M, Powell S, Lattke J, O'Donnell S (2011) Predation and patchiness in the tropical litter: do swarm-raiding army ants skim the

cream or drain the bottle? Journal of Animal Ecology 80: 818–823. https://doi.org/10.1111/j.1365-2656.2011.01826.x

Lach L, Parr C, Abbott K [Eds] (2010) Ant Ecology. Oxford University Press, Oxford. https://doi.org/10.1093/acprof:oso/9780199544639.001.0001

Lamarre GPA, Fayle TM, Segar ST, Laird-Hopkins B, Nakamura A, Souto-Vilarós D, Watanabe S, Basset Y (2020) Monitoring tropical insects in the 21st century. Advances in Ecological Research 62: 295–330. https://doi.org/10.1016/bs.aecr.2020.01.004

Lamb PD, Hunter E, Pinnegar JK, Creer S, Davies RG, Taylor MI (2019) How quantitative is metabarcoding: A meta-analytical approach. Molecular Ecology 28: 420–430. https://doi.org/10.1111/mec.14920

Larsson J (2020) eulerr: area-proportional Euler and Venn diagrams with ellipses. R package version 6.1.0. https://cran.r-project.org/package=eulerr

Leather SR (2017) "Ecological Armageddon"-more evidence for the drastic decline in insect numbers. Annals of Applied Biology 172: 1–3. https://doi.org/10.1111/aab.12410

Leray M, Knowlton N (2015) DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. Proceedings of the National Academy of Sciences 112: 2076–2081. https://doi.org/10.1073/pnas.1424997112

Lister BC, Garcia A (2018) Climate-driven declines in arthropod abundance restructure a rainforest food web. Proceedings of the National Academy of Sciences 115: E10397–E10406. https://doi.org/10.1073/pnas.1722477115

Lucas M, Forero D, Basset Y (2016) Diversity and recent population trends of assassin bugs (Hemiptera: Reduviidae) on Barro Colorado Island, Panama. Insect Conservation and Diversity 9: 546–558. https://doi.org/10.1111/icad.12191

Macfadyen A (1953) Notes on methods for the extraction of small soil arthropods. Journal of Animal Ecology 22: 65–77. https://doi.org/10.2307/1691

McDermott Long O, Warren R, Price J, Brereton TM, Botham MS, Franco AM (2017) Sensitivity of UK butterflies to local climatic extremes: which life stages are most at risk? Journal of Animal Ecology 86: 108–116. https://doi.org/10.1111/1365-2656.12594

Oksanen J, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, Blanchet FG, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H (2018) vegan: community Ecology Package. R package version 2.4-6. https://CRAN.R-project.org/package=vegan

Oliverio AM, Gan H, Wickings K, Fierer N (2018) A DNA metabarcoding approach to characterize soil arthropod communities. Soil Biology and Biochemistry 125: 37–43. https://doi.org/10.1016/j.soilbio.2018.06.026

Palacios-Vargas JG, Castaño-Meneses G, Gómez-Anaya JA (1998) Collembola from the canopy of a Mexican tropical deciduous forest. Pan-Pacific Entomologist 74: 47–54.

Palacios-Vargas JG, Gómez-Anaya JA (1993) Los Collembola (Hexapoda: Apterygota) de Chamela, Jalisco, México (Distribución Ecológica y claves). Folia Entomológica Mexicana 89: 1–34.

Perez KH, Sones JE, deWaard JR, Hebert PDN (2017) Mapping terrestrial biodiversity across the planet: a progress report on the Global Malaise Program. Genome 60: e983. https://doi.org/10.1139/gen-2017-0178

Pimentel D, Andow DA (1984) Pest management and pesticide impacts. International Journal of Tropical Insect Science 5: 141–149. https://doi.org/10.1017/S1742758400008201

Ratnasingham S (2019) mBRAVE: The Multiplex Barcode Research And Visualization Environment. Biodiversity Information Science and Standards 3: e37986. https://doi.org/10.3897/biss.3.37986

Ratnasingham S, Hebert PDN (2013) A DNA-based registry for all animal species: the Barcode Index Number (BIN) system. PLoS ONE 8: e66213. https://doi.org/10.1371/journal.pone.0066213

Roisin Y, Leponce M (2004) Characterizing termite assemblages in fragmented forests: a test case in the Argentinian Chaco. Austral Ecology 29: 637–646. https://doi.org/10.1111/j.1442-9993.2004.01403.x

Römbke J, Sousa JP, Schouten T, Riepert F (2006) Monitoring of soil organisms: a set of standardized field methods proposed by ISO. European Journal of Soil Biology 42: S61–S64. https://doi.org/10.1016/j.ejsobi.2006.07.016

Saitoh S, Aoyama H, Fujii S, Sunagawa H, Nagahama H, Akutsu M, Shinzato N, Kaneko N, Nakamori T (2016) A quantitative protocol for DNA metabarcoding of springtails (Collembola). Genome 59: 705–723. https://doi.org/10.1139/gen-2015-0228

Sánchez-Bayo F, Wyckhuys KA (2019) Worldwide decline of the entomofauna: A review of its drivers. Biological Conservation 232: 8–27. https://doi.org/10.1016/j.biocon.2019.01.020

Shendure J, Ji H (2008) Next-generation DNA sequencing. Nature biotechnology 26: 1135–1145. https://doi.org/10.1038/nbt1486

Smith MA, Janzen DH, Hallwachs W, Longino JT (2015) Observations of *Adelomyrmex* (Hymenoptera: Formicidae) reproductive biology facilitated by digital field microscopy and DNA barcoding. The Canadian Entomologist 147: 611–616. https://doi.org/10.4039/tce.2014.82

Snyder TE (1926) Five new termites from Panama and Costa Rica. Proceedings of the Entomological Society of Washington 28: 7–16.

Southwood TRE, Henderson PA (2000) Ecological Methods. Blackwell Publishing, Oxford.

Stork NE (2018) How many species of insects and other terrestrial arthropods are there on Earth? Annual Review of Entomology 63: 31–45. https://doi.org/10.1146/annurev-ento-020117-043348

Tang M, Hardman CJ, Ji Y, Meng G, Liu S, Tan M, Yang S, Moss ED, Wang J, C, Bruce C, Nevard T, Potts SG, Zhou X, Yu DW (2015) High-throughput monitoring of wild bee diversity and abundance via mitogenomics. Methods in Ecology and Evolution 6: 1034–1043. https://doi.org/10.1111/2041-210X.12416

Townes H (1972) A light-weight Malaise trap. Entomological News 83: 239–247.

van Klink R, Bowler DE, Gongalsky KB, Swengel AB, Gentile A, Chase JM (2020) Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. Science 368: 417–420. https://doi.org/10.1126/science.aax9931

Wagner DL (2020) Insect declines in the Anthropocene. Annual Review of Entomology 65: 457–480. https://doi.org/10.1146/annurev-ento-011019-025151

Wheeler WM (1925) A new guest-ant and other new Formicidae from Barro Colorado Island, Panama. The Biological Bulletin 49: 150–181. https://doi.org/10.2307/1536460

Yang C, Wang X, Miller JA, de Blécourt M, Ji Y, Yang C, Yang C, Wang X, Miller JA, de Blécourt M, Ji Y, Yang C, Harrison RD, Yu DW(2014) Using metabarcoding to ask if easily collected soil and leaf-litter samples can be used as a general biodiversity indicator. Ecological Indicators 46: 379–389. https://doi.org/10.1016/j.ecolind.2014.06.028

Zinger L, Taberlet P, Schimann H, Bonin A, Boyer F, De Barba M, Gaucher P, Gielly L, Giguet-Covex C, Iribar A, Réjou-Méchain M, Rayé

G, Rioux D, Schilling V, Tymen B, Viers J, Zouiten C, Thuiller W, Coissac E, Chave J (2019) Body size determines soil community assembly in a tropical forest. Molecular Ecology 28: 528–543. https://doi.org/10.1111/mec.14919

Supplementary material 1

Methodological considerations for monitoring soil/litter arthropods in tropical rainforests using DNA metabarcoding, with a special emphasis on ants, springtails and termites

Authors: Yves Basset, David A. Donoso, Mehrdad Hajibabaei, Michael T. G. Wright, Kate H. J. Perez, Greg P. A. Lamarre, Luis F. De León, José G. Palacios-Vargas, Gabriela Castaño Meneses, Marleny Rivera, Filonila Perez, Ricardo Bobadilla, Yacksecari Lopez, José Alejandro Ramirez, Héctor Barrios

Data type: Appendices, protocols, occurrences, supplementary tables and figures

Explanation note: **Appendix S1.** Suggested field procedures for collecting soil/litter arthropods in a tropical rainforest for long-term monitoring with DNA metabarcoding. **Table S1.** Main characteristics of laboratory and bioinformatic procedures used for the three datasets covered by this study. **Fig. S1.** Euler diagrams for (a) soil ants, (b) arboreal ants, (c) soil termites and (d) arboreal termites, indicating the number of species detected by metabarcoding and by each collecting method. **Fig. S2.** Euler diagrams for common species of (a) soil ants, (b) springtails, and (c) soil termites, indicating the number of species detected by metabarcoding and by each collecting method. **Fig. S3.** Accumulation of species richness vs. the number of samples for common species nesting or dwelling in the soil of (a) ants, (b) springtails and (c) termites, detailed for each collecting method: blue = soil samples; green = Malaise trap samples; red = Light trap samples.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/mbmg.4.58572.suppl1

Supplementary material 2

Appendix S2

Authors: Yves Basset, David A. Donoso, Mehrdad Hajibabaei, Michael T. G. Wright, Kate H. J. Perez, Greg P. A. Lamarre, Luis F. De León, José G. Palacios-Vargas, Gabriela Castaño Meneses, Marleny Rivera, Filonila Perez, Ricardo Bobadilla, Yacksecari Lopez, José Alejandro Ramirez, Héctor Barrios

Data type: table

Explanation note: Species occurrence in soil, Malaise and light trap samples.

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/mbmg.4.58572.suppl2